استخراج ویژگی‌های مکان- فرکانسی جهت بازیابی تصویر چهره از پایگاه داده حجیم تصاویر

محمد مهدی بخشی[۱]، دانشجوی کارشناسی ارشد؛ حمید حسن پور[۲]، استاد؛ منصور فاتح[۳]، استادیار

۱-دانشکده مهندسی کامپیوتر و فناوری اطلاعات- دانشگاه صنعتی شاهرود- شاهرود- ایران - mahdi.bakhshi68@hotmail.com
۲- دانشکده مهندسی کامپیوتر و فناوری اطلاعات- دانشگاه صنعتی شاهرود- شاهرود- ایران - h.hassanpour@shahroodut.ac.ir
۳- دانشکده مهندسی کامپیوتر و فناوری اطلاعات- دانشگاه صنعتی شاهرود- شاهرود- ایران - mansoor_fateh@shahroodut.ac.ir

چکیده: در این مقاله، روشی نوین جهت استخراج ویژگی از تصاویر چهره ارائه شده است. هدف اصلی، بازیابی چهره از پایگاه داده حجیم است. با افزایش حجم پایگاه داده، تفکیک کردن چهره‌های مختلف با مشکل مواجه می‌شود. روش پیشنهادی با استخراج ویژگی‌های مناسب، فاصله بین چهره‌های افراد در فضای ویژگی را افزایش می‌دهد. این روش بر اساس عملکرد سامانه بینایی انسان است و به صورت ترتیبی از کل به جزء، ویژگی استخراج می‌کند. برای این منظور، از ویژگی‌های مکان- فرکانسی استفاده شده است. در این روش، با اعمال پنجره‌های هم مرکز با ابعاد مختلف روی تصویر چهره، محتوای هر پنجره به فضای فرکانسی منتقل می‌شود. تغییر مؤلفه‌های فرکانسی در پنجره‌های مختلف، فضای ویژگی تصویر را تشکیل می‌دهند. سپس، با استفاده از فیلتر مناسب، تنها مولفه‌های فرکانسی با قابلیت بالا در جداسازی بین تصاویر چهره افراد مختلف، حفظ می‌شوند. در نهایت، با استفاده از معیار فاصله‌ی اقلیدسی، تصویر نهایی از پایگاه داده بازیابی می‌شود. در این مقاله از پایگاه داده FERET استفاده شده است. نرخ بازشناسی در مقایسه با بهترین روش قبلی با ویژگی‌های مشابه روش پیشنهادی، در حجم مشابه، با ۲٪ افزایش به ۹۹٪ ارتقا پیدا کرده است. با افزایش حجم پایگاه داده به ۹۹۰ چهره متمایز، نرخ بازشناسی ۹۰/۴٪ حاصل شده است.

واژه‌های کلیدی: بازشناسی چهره، استخراج ویژگی، حوزه مکان- فرکانسی، پایگاه داده حجیم، انتخاب ویژگی.

Spatial-Frequency Features Extracting for Facial Image Retrieval from a Big Image Database

Mohammad. M. Bakhshi[1], MSc student; Hamid Hassanpour[2], Professor; Mansoor Fateh[3], Assistant Professor

1- Faculty of Computer and IT Engineering, Shahrood University of Technology, Shahrood, Iran, Email: mahdi.bakhshi68@hotmail.com
2- Faculty of Computer and IT Engineering, Shahrood University of Technology, Shahrood, Iran, Email: h.hassanpour@shahroodut.ac.ir
3- Faculty of Computer and IT Engineering, Shahrood University of Technology, Shahrood, Iran, Email: mansoor_fateh@shahroodut.ac.ir

Abstract: In this paper, a new method is presented to feature extraction from facial images. The main purpose of this paper is probe image retrieval from a big database. By increasing the size of the database, the similarities between people increases and the separation capability decreases. The proposed method increases the distance between peoples in feature space by extracting appropriate features. This method is based on properties of the human vision system and sequentially extracts features in top-down manner. For this purpose, spatial- frequency features are used. In this method, by applying concentric windows in different size on the facial image, the content of each window are mapped to frequency space. The change of frequency components in different windows forms the feature space of image. Then frequency component with high separation capability between face images is remained by appropriate filter. In the end, the final image is retrieved from database by Euclidean distance criterion. In this paper the FERET database is used. Recognition rate compared with the best current method in similar size of database, with 2% improvement reached to 99%. By increasing database size to 990 classes, 90.4% of recognition rate is achieved.

Keywords: Face recognition, Feature extraction, Spatial-Frequency domain, Big database, Feature selection.

تاریخ ارسال مقاله: ۱۳۹۵/۱۰/۰۵
تاریخ اصلاح مقاله: ۱۳۹۵/۱۱/۱۳، ۱۳۹۶/۰۲/۰۹ و ۱۳۹۶/۰۵/۲۲
تاریخ پذیرش مقاله: ۱۳۹۶/۰۸/۱۱
نام نویسنده مسئول: منصور فاتح
نشانی نویسنده مسئول: ایران - شاهرود - بلوار دانشگاه - دانشگاه صنعتی شاهرود- دانشکده مهندسی کامپیوتر و فناوری اطلاعات.

۱- مقدمه

بازشناسی چهره در یک دهه‌ی اخیر، بخش عظیمی از تحقیقات حوزه بازشناسی الگو و بینایی ماشین را به خود اختصاص داده است. بازشناسی چهره در انسان‌ها با دقت بالایی انجام می‌شود درحالی که بازشناسی هوشمند چهره دشوار است. روش‌های زیادی برای ارتقای دقت بازشناسی هوشمند چهره ارائه شده است. با توجه به ارائه روش‌های گوناگون در این حیطه، این تکنولوژی به شکل قابل توجهی پیشرفت کرده است[۱]. روش‌های بازشناسی هوشمند چهره را می‌توان به سه دسته تقسیم نمود: ۱) روش‌های ظاهرگرا[۱]، ۲) روش‌های مدل‌گرا[۲] ، ۳) روش‌های ترکیبی[۳] که از هر دو روش مذکور تواماً استفاده می‌کنند[۲]. روش‌های ظاهرگرا، ویژگی‌های خود را از تمام محدوده چهره استخراج می‌کنند. این دسته از روش‌ها در صورت بهره‌مندی از روش‌های مناسب مبتنی بر استخراج ویژگی، می‌توانند اطلاعات مفید تمام چهره را استخراج نمایند. اندازه‌ی بردار ویژگی در این گروه از روش‌ها، زیاد است. از این رو، این روش‌ها باید با کاهش ویژگی همراه شوند که منجر به از دست دادن برخی ویژگی‌های مهم و کاهش دقت بازیابی می‌شوند.

روش‌های مدل‌گرا با تمرکز بر استخراج ویژگی از نواحی خاص چهره، فرایند بازشناسی را انجام می‌دهند. این روش‌ها معمولاً در حجم کم پایگاه داده، قدرت بازشناسی مناسبی دارند. اما با افزایش حجم پایگاه داده و ایجاد شباهت در بردارهای ویژگی، دقت این روش‌ها کاهش می‌یابند. روش‌های ترکیبی، ترکیبی از دو روش ظاهرگرا و مدل‌گرا هستند. این روش‌ها، از فرآیند استخراج ویژگی‌ها در تمام چهره استفاده می‌کنند. این روش‌ها از روش‌های ظاهرگرا سریع‌تر و از روش‌های مدل‌گرا دقیق‌ترند[۳].

بازشناسی هوشمند چهره انسان همواره با مشکل‌هایی نظیر حجم زیاد چهره‌های موجود در پایگاه داده و در نتیجه تشابه ویژگی‌های استخراجی و کاهش سرعت بازیابی، لحاظ نکردن تغییر حالت‌های چهره[۴]، تغییر روشنایی و زاویه دید در هنگام بازیابی، وجود تنها یک تصویر به ازای هر شخص در زمان یادگیری و بسیاری مشکلات دیگر مواجه است. در پژوهش‌های متعدد، تلاش‌هایی جهت رفع هر یک از این معایب انجام گرفته است[۳].

ابتدایی‌ترین روش‌های بازشناسی چهره به عنوان پایه‌ی روش‌های جدید، استفاده می‌شوند. روش چهره‌های ویژه[۴] از جمله این روش‌ها است[۵]. این روش با محاسبه میانگین و ماتریس کواریانس تصاویر، فضای ویژگی را از تصاویر استخراج می‌کند. در مراجع [۶، ۷] روش الگوی باینری محلی جهت بازشناسی چهره ارائه شده است و در [۸]، روشی مبتنی بر یادگیری ماشین برای بازشناسی چهره پیشنهاد شده است. چن[۵] روشی برای بازشناسی چهره با سرعت بالا ارائه کرده است[۹]. این روش، مبتنی بر تبدیل کسینوسی گسسته و تحلیل خطی فیشر[۶] و شبکه عصبی RBF[۷] است. در مرجع [۱۰] الگوریتم SIFT[۸] در بازشناسی چهره به کار گرفته شده است. این الگوریتم مبتنی بر محاسبه دو پارامتر زاویه و اندازه از نقاط کلیدی تصویر است. این نقاط، کوچک‌ترین و بزرگترین نقاط هرم گاوسی می‌باشند که از اعمال تابع گاوسی در پهناهای متفاوت روی تصویر تشکیل می‌شود. در این پژوهش حجم پایگاه داده انتخابی نسبت به روش‌های دیگر بیشتر است. سپس ترکیبی از دو الگوریتم SIFT و SURF[۹] در [۱۱] ارائه شده است که دقت بالاتری در بازشناسی چهره به دست آورده است، هرچند به حجم دقیق پایگاه داده در این روش اشاره نشده است. ژو[۱۰] و همکاران در [۱۲] با روش استخراج ویژگی موجک گابور و همینطور انتخاب ویژگی توسط الگوریتم آدابوست[۱۱]، توانستند نتایج نسبتاً مطلوبی را حاصل کنند. در [۱۳] روشی ترکیبی با استفاده از الگوریتم PCA[۱۲] و موجک گابور ارائه شده است.

در [۱۴] الگوریتمی مبتنی بر تبدیل هاف جهت فائق آمدن بر مشکل تغییر روشنایی چهره ارائه شده است. همچنین در [۱۵] دو روش پیش پردازش ارائه شده و سپس از فیلتر گابور جهت مرتب‌سازی تصاویر چهره در راستای تغییرات زاویه استفاده شده است. در [۱۶] روشی ترکیبی جهت بازشناسی چهره ارائه شده است. در این روش، ابتدا با افزایش فضای ویژگی و سپس با کاهش بُعد همراه با حفظ اطلاعات، ویژگی‌های تصویر استخراج شده‌اند. در [۱۷] با استفاده از ماتریس آنتروپی، کواریانس، PCA و در نهایت تعریف ماتریس فاصله آنتروپی، عملیات بازشناسی چهره انجام شده است. در [۱۸] با استفاده از روش بازسازی محلی مبتنی بر یک چهره واحد، تصاویر را به صورت بلاکی پیمایش کرده و با مقایسه بلاک‌های بازسازی شده‌ی متناظر بین دو تصویر، تشابه تصاویر اندازه‌گیری شده‌اند. در [۲۱-۱۹] نیز روش‌هایی برای بازیابی چهره با استفاده از تشابه نقاط مکانی و ساختاری، شبکه عصبی RBF و دسته بند C- میانگین فازی[۱۳] ارائه شده‌اند.

در بیشتر روش‌های یاد شده که تنها از یک تصویر به ازای هر شخص استفاده می‌کنند، از تعداد افراد کمتر از ۵۰ نفر در پایگاه داده جهت ارزیابی استفاده شده است. همچنین در روش‌هایی با پایگاه داده‌های حجیم، چندین تصویر به ازای هر شخص جهت فرآیند آموزش به کار گرفته شده است. لذا مقایسه‌ی روش پیشنهادی تنها با بهترین روش‌ها با شرایط مشابه انجام شده است. این شرایط شامل استفاده‌ی تنها یک تصویر به ازای هر شخص و عدم نیاز به آموزش الگوریتم است. عدم نیاز به آموزش، موجب افزایش کاربرد، بهبود سرعت و کاهش حجم ذخیره‌سازی داده‌ها می‌شود. در روش پیشنهادی، تعداد تصاویر پایگاه داده به مراتب بیشتر از روش‌های پیشین بوده و در پایگاه داده به ازای هر شخص تنها یک تصویر لحاظ شده است. بنابراین مشکل پایگاه داده حجیم و وجود تنها یک تصویر به ازای هر شخص، تا حد قابل توجهی رفع شده است. همچنین مرحله آزمون با استفاده از تنها یک تصویر به ازای هر شخص انجام می‌شود. نوآوری اصلی روش پیشنهادی، ارائه‌ی روشی جدید در استخراج ویژگی و بازیابی تصویر چهره در پایگاه داده‌های حجیم است. یکی از دلایل عدم رفع این مشکل در روش‌های پیشین، عدم تفکیک پذیری ویژگی‌های افراد مختلف در حجم پایگاه داده انبوه است. با افزایش حجم پایگاه داده، قابلیت تفکیک ویژگی‌های افراد

مختلف کاهش پیدا می‌کند. در روش پیشنهادی، با استخراج ویژگی‌های مناسب، فاصله‌ی بین تصاویر افراد مختلف در فضای ویژگی افزایش یافته و قابلیت جداسازی افراد بهبود می‌یابد.

تلاش اصلی در سامانه‌های بازشناسی چهره، نزدیک کردن ویژگی‌های استخراج شده به ویژگی‌های بصری انسان است. روش پیشنهادی نیز در راستای این هدف طراحی شده است. ایده اصلی این روش مبتنی بر خاصیت بازشناسی اشخاص توسط قوه بصری انسان است. انسان‌ها هنگام بازشناسی شخص خاص در بین انبوهی از افراد، با پیمایش ترتیبی تصویر از کل به جزء و بررسی چهره در هر مرحله، به بازشناسی شخص می‌پردازند. در هر مرحله از پیمایش، استخراج معنایی تصویر چهره، انجام می‌شود. در ادامه، ویژگی‌های تصویر مورد نظر با ویژگی‌های موجود در ذهن، مقایسه می‌شود و جهت اطمینان، جستجو ادامه می‌یابد.

الگوریتم پیشنهادی نیز با استفاده از این منطق طبیعی، بازشناسی چهره‌ی انسان را انجام داده است. برای استفاده از این منطق در ماشین‌های بازشناسی هوشمند چهره، از اصول ریاضیات کاربردی استفاده شده است.

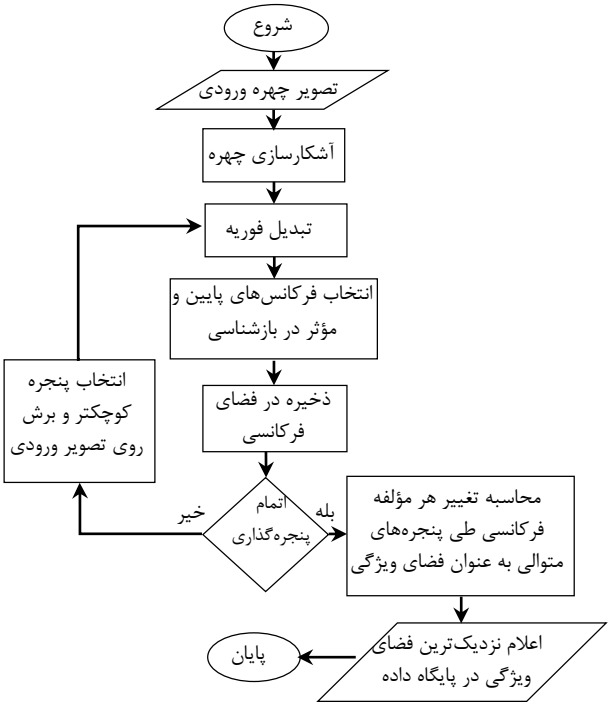
روش پیشنهادی دو بخش اصلی آشکارسازی چهره و بازشناسی چهره دارد. این روش در بخش بازشناسی چهره، شامل سه مرحله اصلی است: ۱) استخراج ویژگی از ناحیه چهره هر شخص ۲) کاهش فضای ویژگی ۳) اندازه‌گیری فاصله تصویر آزمون و تصاویر پایگاه داده.

در ادامه، روش پیشنهادی شرح داده شده است. این بخش شامل شرح روش استخراج ویژگی و تعریف پارامترها، کاهش ویژگی‌ها و معیار شباهت در بازشناسی چهره است. دربخش سوم، به بررسی نتایج آزمایشی الگوریتم و مقایسه آن با چند روش دیگر پرداخته شده است. در نهایت در بخش آخر، نتیجه‌گیری ارائه شده است.

۲- روش پیشنهادی

در روش پیشنهادی ابتدا ناحیه‌ی چهره، استخراج می‌شود. تمرکز اصلی این مقاله بر روی بازشناسی چهره است. از این رو، از الگوریتم Viola-Jones و توابع از پیش تعریف شده در نرم‌افزار متلب برای آشکارسازی ناحیه چهره استفاده شده است[۲۲]. در ادامه، از تصویر ورودی طی پنجره‌گذاری‌های متوالی استخراج معنا می‌شود که با استفاده از آن‌ها، فضای ویژگی تصویر ایجاد می‌شود. سپس، با استفاده از معیار فاصله‌ی مناسب، فاصله تصویر آزمون با تصاویر پایگاه داده محاسبه می‌شود. در نهایت، نزدیک‌ترین تصویر به تصویر آزمون، به عنوان جواب نهایی انتخاب می‌شود. همانگونه که ذکر شد، نوآوری اصلی روش پیشنهادی، ارائه‌ی روشی نوین در استخراج ویژگی و بازیابی تصویر چهره در پایگاه داده‌های حجیم است. روش پیشنهادی نیز همانند سایر الگوریتم‌های بازشناسی چهره، قابل اعمال روی تصاویر چهره از رو برو می‌باشد و تصاویر چهره دارای زاویه سر قابل بازشناسی نمی‌باشند. فرایند کلی الگوریتم پیشنهادی در شکل ۱ به تصویر کشیده شده است.

۲-۱- استخراج ویژگی



شکل ۱ : فرآیند کلی الگوریتم ارائه شده برای بازشناسی چهره

بخش اصلی سامانه بازشناسی چهره، استخراج ویژگی از تصویر چهره است. تمرکز اصلی الگوریتم پیشنهادی نیز معطوف به این قسمت است. استخراج ویژگی در این بخش، به صورت اعمال ترتیبی از پنجره‌های هم مرکز با اندازه‌های متفاوت روی تصویر چهره انجام می‌شود. این ویژگی‌ها مبتنی بر تبدیل فوریه دو بعدی هستند. در هر مرحله، تصویر را از حاشیه به سمت مرکز تصویر پنجره‌گذاری می‌کنیم. نتیجه این فرایند در مرحله pام پنجره‌گذاری، به شکل ریاضی در رابطه (۱) نشان داده شده است. در این رابطه، I تصویر چهره‌ی ورودی با ابعاد L×L و h^p تصویر حاصل از پنجره‌گذاری در مرحله pام است. در این رابطه همواره $2 < N < L$ بوده و در هر مرحله N=L-2(p-1) می‌باشد.

$$h^{p}_{N\times N}(i,j) = I(p+i-1, p+j-1) \qquad (۱)$$

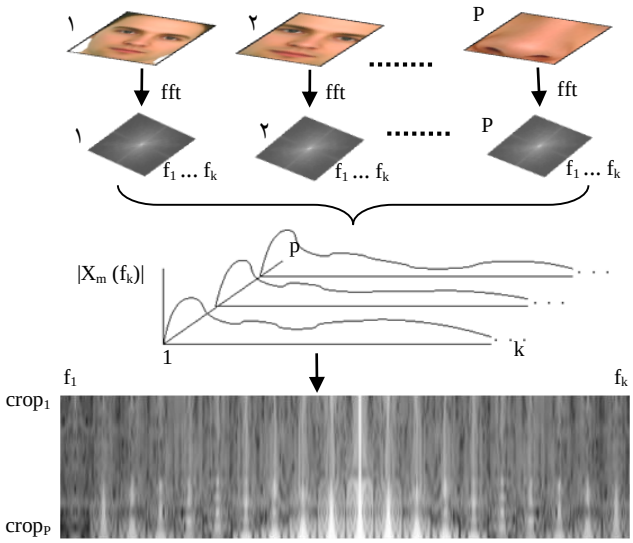
میزان تغییر اندازه پنجره‌ها، متناسب با تعداد پیکسل‌های تعیین شده توسط کاربر است. با اعمال تبدیل فوریه دوبعدی روی تصویر حاصل از هر مرحله پنجره‌گذاری، اندازه‌ی مولفه‌های فرکانسی استخراج می‌شوند. این تبدیل در رابطه (۲) نشان داده شده است. در این رابطه، h تصویر ورودی و H تبدیل فوریه تصویر است. n و m معرف موقعیت پیکسل تصویر ورودی، k و l بیانگر موقعیت مؤلفه فرکانسی تصویر ورودی در صفحه تبدیل یافته فرکانسی هستند.

$$H(\mathrm{k},\mathrm{l}) = \sum_{\mathrm{n}=0}^{\mathrm{M}-1}\sum_{\mathrm{m}=0}^{\mathrm{M}-1} \mathrm{e}^{-2\pi i\left(\frac{kn}{M}+\frac{\mathrm{lm}}{M}\right)} \mathrm{h}(\mathrm{n},\mathrm{m}) \qquad (۲)$$

ویژگی‌های نهایی مبتنی بر تغییرات اندازه‌ی هر مؤلفه فرکانسی طی پنجره‌های متوالی، هستند. ویژگی‌های نهایی جهت بهبود مقیاس و قابلیت مقایسه دقیق‌تر، به فضای دسیبل[۱۴] انتقال می‌یابند. این تغییرات به عنوان فضای ویژگی چهره‌ی هر شخص معرفی می‌شوند. منظور از مؤلفه‌های فرکانسی، پیکسل‌های تبدیل یافته‌ی تصویر در فضای فوریه دو بعدی تصویر هستند.

دلایل استفاده از ویژگی‌های فوریه، قدرت بالای استخراج خواص معنایی تصاویر، ثابت بودن نسبت به چرخش تصویر و سرعت قابل قبول آن است. در بازشناسی چهره، سرعت مناسب و مقاومت نسبت به چرخش از ویژگی‌های بسیار مهم بشمار می‌رود.

در سیگنال‌های یک بعدی، تبدیل فوریه زمان کوتاه[۱۵] جهت استخراج معنایی در کاربردهای فراوانی استفاده می‌شود که مشابه فرآیند فوق است. از جمله این کاربردها، می‌توان به درک معنایی گفتار اشاره کرد[۲۳]. در این مقاله، ویژگی‌های تغییرات فرکانسی در راستای مکان، در بازشناسی چهره استفاده شده‌اند که بیانگر مفاهیمی بسیار سودمند در تصاویر هستند. در شکل ۲ فرایند استخراج ویژگی نشان داده شده و ماتریس ویژگی یک تصویر نمونه استخراج شده است.



شکل ۲: فرآیند استخراج ویژگی از هر تصویر چهره.
P: تعداد پنجره‌گذاری. k: تعداد مولفه‌های فرکانسی انتخاب شده. X_m: تبدیل یافته تصویر در حوزه فرکانس.

روش پیشنهادی شامل پارامترهایی می‌باشد که در زیر به تعریف آن‌ها پرداخته‌ایم:

۱) ابعاد تصویر ورودی چهره: L×L.

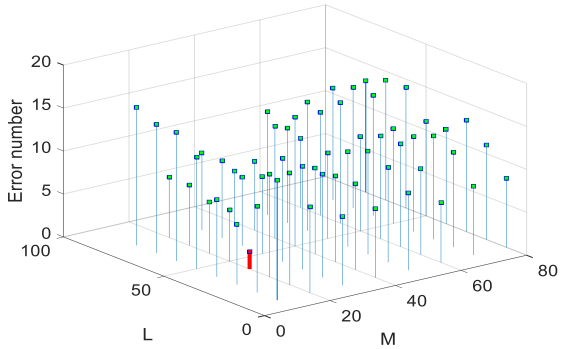
۲) در هر مرحله، ابعاد پنجره اعمال شده روی تصویر چهره، به اندازه‌ی تعداد پیکسلی مشخص کاهش پیدا می‌کند و این فرایند به سمت مرکز تصویر ادامه پیدا می‌کند. تعداد پیکسل مذکور توسط μ مشخص می‌شود.

۳) ابعاد تبدیل فوریه دوبعدی مورد استفاده در تمامی مراحل الگوریتم توسط M×M نمایش داده شده است. به عبارتی M^2 تعداد مولفه‌های فرکانسی تعریف شده است.

تعداد دفعات پنجره‌گذاری روی تصویر از حاشیه به سمت مرکز تصویر، با استفاده از پارامترهای فوق طبق رابطه (۳) محاسبه می‌شود. با افزایش پارامتر μ، اندازه ماتریس ویژگی‌ها کاهش یافته و سرعت بازشناسی چهره افزایش می‌یابد. با کاهش این پارامتر، دقت الگوریتم افزون می‌شود. در دقیق‌ترین حالت، مقدار μ برابر با ۱ است.

$$P = \frac{\frac{L}{2} - \mu}{\mu} \qquad (۳)$$

تغییر دو پارامتر M و L نیز، نرخ بازشناسی و سرعت بازیابی تصویر را دستخوش تغییر می‌کنند. افزایش این دو پارامتر، فضای ویژگی را افزون و سرعت بازیابی را کاهش می‌دهند. لذا باید مقداری بهینه برای این دو پارامتر انتخاب شود تا به ازای آن‌ها نرخ بازیابی بیشینه شود. مقدار بهینه این دو پارامتر طی آزمایش‌های متوالی روی پایگاه داده شامل ۲۰۰ چهره تصادفی، محاسبه شدند. نتایج برخی از این آزمایش‌ها در محدوده نقاط مشخص در شکل ۳ نمایش داده شده است. چنانچه در شکل ۳ مشاهده می‌شود، مقادیر بهینه‌ی یافت شده برای این دو پارامتر، L=۴۰ و M=۲۰ است.



شکل ۳ : مقادیر بهینه دو پارامتر M و L در روش ارائه شده

چنانچه مشخص است، با افزایش یا کاهش بیش از حد این پارامترها، نرخ بازشناسی به مراتب کاهش می‌یابد. لذا محدوده کمتر از ۸۰ و بیشتر از ۱۰ نمایش داده نشده‌اند. در ادامه‌ی مقاله، از این مقادیر بهینه در آزمایش‌ها استفاده شده است.

۲-۲- کاهش ویژگی‌ها

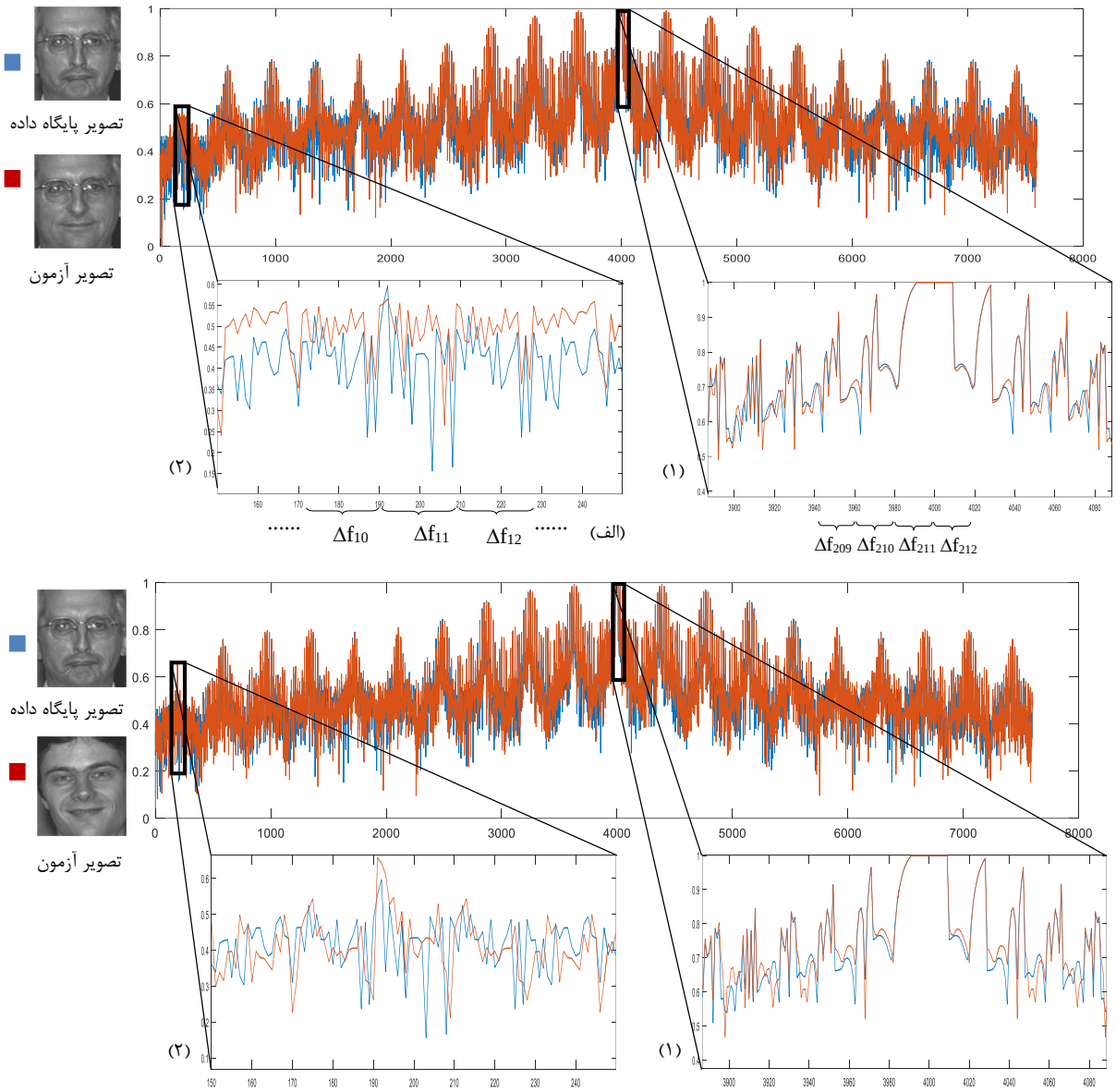
پس از استخراج ویژگی، نیاز به انتخاب ویژگی‌های سودمند و صرف نظر کردن از ویژگی‌های غیرمفید است. در روش پیشنهادی نیز جهت کاهش ابعاد ویژگی، افزایش سرعت و همینطور افزایش دقت بازیابی، فضای ویژگی کاهش داده شده است. به همین منظور مولفه‌های فرکانسی غیر سودمند استخراج شده از تصویر حذف می‌شوند تا علاوه بر افزایش دقت بازیابی، سرعت بازیابی افزایش یابد. برای این منظور، بررسی‌های فراوانی بر روی فضای ویژگی تصویر یک شخص در حالت‌های مختلف و فضای ویژگی دو شخص متفاوت انجام شد. در این بررسی‌ها مشخص شد که در روش پیشنهادی، فرکانس‌های پایین اهمیت بیشتری نسبت به فرکانس‌های بالا دارند. فرکانس‌های بالا نشان‌دهنده‌ی بافت تصویر

می‌باشند که در انسان‌های مختلف به هم نزدیک هستند. اما فرکانس‌های پایین وابسته به موقعیت و اندازه‌ی اجزاء تشکیل دهنده‌ی چهره نظیر بینی، لب و چشم هستند که در افراد مختلف متفاوتند.

در شکل ۴ ماتریس ویژگی‌های استخراج شده برای افراد نمونه، به صورت یک بردار واحد به تصویر کشیده شده است. همانگونه که مشاهده

فرکانس‌های بالا موجب بازشناسی ناصحیح می‌شوند. لذا برای افزایش نرخ بازیابی در روش پیشنهادی باید ویژگی‌های فرکانس پایین انتخاب شوند. انتخاب این مولفه‌های فرکانسی با اعمال فیلتری بر روی تمام فضای فرکانسی، میسر می‌شود.

فیلترهای حوزه‌ی فرکانس که تاکنون در تصاویر استفاده شده است،

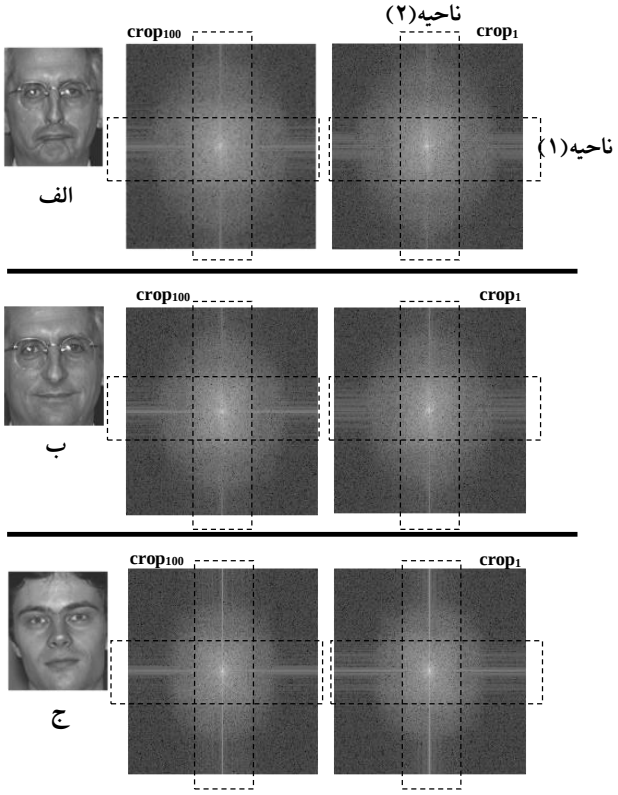


شکل ۴: نمایش فضای ویژگی تصاویر چهره نمونه به صورت برداری (با پیمایش ستونی ماتریس ویژگی‌های استخراج شده). (الف) مقایسه دو تصویر از یک شخص. (ب) مقایسه دو تصویر از دو شخص متفاوت. Δf_i نشان دهنده تغییرات مؤلفه فرکانسی i ام در پنجره‌های متوالی تصویر است.

می‌شود، تغییرات مولفه‌های فرکانس پایین در دو تصویر از یک شخص، مشابه (شکل ۴-الف-۱) و در دو تصویر از دو شخص متفاوت مغایر هم هستند (شکل ۴-ب-۱). اما تغییرات مولفه‌های فرکانس بالا در دو تصویراز یک شخص، مغایر هم هستند (شکل ۴-الف-۲). به عبارت دیگر در روش پیشنهادی، فرکانس‌های پایین موجب بازشناسی صحیح و اکثر

تصویر را از حالتی به حالت دیگر تغییر می‌دهند و جهت مقایسه یک تصویر با خودش پس از تغییرات، استفاده می‌شوند. اما در این تحقیق، فیلتر باید به گونه‌ای طراحی شود که ویژگی‌های منحصر به فرد تصویر ورودی، جهت مقایسه با تصاویر دیگر حفظ شوند. از این‌رو، فیلترهای حوزه فرکانس متداول، در این کاربرد کارا نیستند. برای منحصربه‌فرد

کردن تصاویر باید طی پنجره‌گذاری‌های متوالی، نواحی پرتغییر تصاویر در فضای فرکانسی، استخراج و نواحی کم‌تغییر فیلتر شوند. طی آزمایش‌های فراوان، نواحی حساس و تغییرپذیر در فضای فوریه تصویر چهره مشخص شدند. این نواحی در شکل ۵ نشان داده شده است. سایر قسمت‌های تصویر دارای اطلاعات غیر ضروری در بازشناسی چهره هستند و باعث بازشناسی کاذب می‌شوند. بنابراین در گام اول نواحی غیر ضروری فیلتر می‌شوند و در گام دوم فیلتری مناسب نیز برای نواحی ضروری طراحی می‌شود.



شکل ۵ : مولفه‌های فرکانسی در دو برش مشخص از سه تصویر نمونه (تغییرات اصلی در راستای برش‌ها، در دو ناحیه دسته بندی شده‌اند). الف: تصویر پایگاه داده. ب: تصویر آزمون همنوع. ج: تصویر آزمون غیر همنوع

در شکل ۵ فضای فوریه سه تصویر در نواحی ضروری، یعنی دو ناحیه (۱) و (۲) در فضای فوریه، مورد مقایسه قرار گرفته‌اند. چنانچه در شکل ۵ مشاهده می‌شود، تغییرات مولفه‌های ناحیه (۱) طی برش ۱ تا برش ۱۰۰ در هر سه تصویر، ناهمگون و غیر قابل تفکیک هستند. در حالیکه تغییرات مولفه‌های ناحیه (۲) برای دو تصویر الف و ب از یک شخص، از تصویر ج قابل تفکیک هستند. همانگونه که اشاره شد، محدوده‌ای از فرکانس حائز اهمیت است که تغییرات آن طی پنجره‌گذاری‌های متوالی بین دو تصویر از یک شخص و دو شخص متفاوت قابل تفکیک باشد. لذا

$$distance\left(I_1,I_2\right)=\frac{\sum_{i=1}^{k}\sqrt{\sum_{j=1}^{P}\left(I1_{ij}-I2_{ij}\right)^2}}{k} \qquad (۴)$$

اهمیت ناحیه (۲) به مراتب بیشتر از ناحیه (۱) است. بنابراین در فیلتر نهایی، اهمیت ناحیه (۲) نیز باید لحاظ شود. ماسک طراحی شده در قالب شکل ۶ خواص مذکور را دارد و در این مقاله برای طراحی فیلتر، از آن استفاده می‌شود. این ماسک به صورت صفر - یک طراحی شده است و پس از اعمال آن در حوزه فرکانس تصویر چهره، تنها مولفه‌های فرکانسی متناظر با پیکسل‌های غیر صفر ماسک، حفظ می‌شوند.

0	0	0	0	1	1	1	1	1	1	1	1	1	1	0	0	0	0
0	0	0	0	1	1	1	1	1	1	1	1	1	1	0	0	0	0
0	0	0	0	1	1	1	1	1	1	1	1	1	1	0	0	0	0
0	0	0	0	1	1	1	1	1	1	1	1	1	1	0	0	0	0
0	0	0	0	1	1	1	1	1	1	1	1	1	1	0	0	0	0
0	0	0	0	1	1	1	1	1	1	1	1	1	1	0	0	0	0
0	0	0	0	1	1	1	1	1	1	1	1	1	1	0	0	0	0
0	0	0	0	1	1	1	1	1	1	1	1	1	1	0	0	0	0
0	0	0	0	1	1	1	1	1	1	1	1	1	1	0	0	0	0
0	0	0	0	1	1	1	1	1	1	1	1	1	1	0	0	0	0
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
0	0	0	0	1	1	1	1	1	1	1	1	1	1	0	0	0	0
0	0	0	0	1	1	1	1	1	1	1	1	1	1	0	0	0	0
0	0	0	0	1	1	1	1	1	1	1	1	1	1	0	0	0	0
0	0	0	0	1	1	1	1	1	1	1	1	1	1	0	0	0	0
0	0	0	0	1	1	1	1	1	1	1	1	1	1	0	0	0	0
0	0	0	0	1	1	1	1	1	1	1	1	1	1	0	0	0	0
0	0	0	0	1	1	1	1	1	1	1	1	1	1	0	0	0	0
0	0	0	0	1	1	1	1	1	1	1	1	1	1	0	0	0	0
0	0	0	0	1	1	1	1	1	1	1	1	1	1	0	0	0	0
0	0	0	0	1	1	1	1	1	1	1	1	1	1	0	0	0	0

M×M

شکل ۶ : ماسک طراحی شده جهت فیلتر فضای فرکانسی تصاویر چهره در روش پیشنهادی.

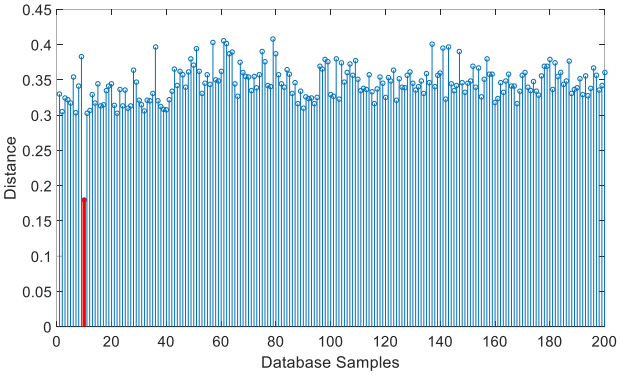
با اعمال این فیلتر بر روی فضای فرکانسی تصویر چهره در هر مرحله از پنجره‌گذاری، فضای فرکانسی ۴۲٪ کاهش پیدا می‌کند. همچنین نرخ بازشناسی چهره نیز به مراتب افزایش می‌یابد که نشان دهنده انتخاب ماسک مناسب جهت فیلتر فضای فرکانسی تصاویر چهره است.

با اعمال فیلتر روی پایگاه داده شامل ۲۰۰ چهره متمایز، نرخ بازشناسی نسبت به حالت اولیه ۲/۵٪ افزایش داشته و زمان بازیابی نیز پس از اعمال فیلتر ms۱۴۳/۱۱ کاهش پیدا کرده است.

۲-۳- معیار فاصله

پس از استخراج فضای ویژگی از هر تصویر چهره، باید فاصله هر چهره با چهره‌های موجود در پایگاه داده محاسبه شود. برای این امر، ابتدا دو فضای ویژگی فرضی I_1 و I_2 برای دو تصویر متفاوت تعریف می‌شود. این فضاهای ویژگی، k مؤلفه فرکانسی دارند و هر مؤلفه فرکانسی، P مقدار طی پنجره‌های متوالی از تصویر دارد. در واقع هر بردار ویژگی، k*P ویژگی دارد. در ادامه، فاصله اقلیدسی بردار تغییرات هر مؤلفه فرکانسی در I_1، با بردار تغییرات مؤلفه فرکانسی متناظر خود در I_2 محاسبه می‌شود. در واقع هر بردار ویژگی، به k دسته P تایی تقسیم می‌شود و فاصله هر دسته، از دسته متناظر در تصویر دوم محاسبه می‌شود. با داشتن k دسته، در نهایت k مقدار حاصل می‌شود. با میانگین‌گیری این k مقدار، فاصله دو تصویر محاسبه می‌شود. به این ترتیب فاصله تصویر آزمون با تمام تصاویر موجود در پایگاه داده محاسبه شده و نزدیک‌ترین فاصله به عنوان محتمل‌ترین جواب در نظر گرفته می‌شود. در رابطه‌ی (۴) این فاصله نمایش داده شده است. در این رابطه، k تعداد مولفه‌های فرکانسی، P تعداد پنجره‌گذاری‌های تصویر و I_1 و I_2 تبدیل یافته دو تصویر به فضای ویژگی در حوزه فرکانس هستند.

آزمایش‌ها نشان می‌دهد که تصویر آزمون، از تصویر هم برچسب خود در پایگاه داده، فاصله اندکی دارد و بازشناسی چهره با دقت بالایی انجام می‌شود. شکل ۷ نمونه‌ای از فاصله یک تصویر آزمون با تمامی تصاویر موجود در پایگاه داده را نمایش می‌دهد. در این مثال حجم پایگاه داده ۲۰۰ تصویر چهره متمایز در نظر گرفته شده است. در این نمونه، فاصله‌ی تصویر آزمون از تصویر هم‌برچسب خود اندک است و با دومین کم‌ترین فاصله، اختلاف قابل توجهی دارد. در واقع، فاصله تصویر آزمون تا تصویر هم برچسب خود در پایگاه داده، به طرز قابل توجهی قابلیت جداسازی از سایر فاصله‌ها را دارد. این خاصیت نشان دهنده‌ی قابلیت تفکیک چهره‌ها در فضای ویژگی با استفاده از روش ارائه شده می‌باشد.



شکل ۷: مقادیر فاصله تصویر آزمون از ۲۰۰ تصویر پایگاه داده. بازشناسی صحیح دهمین داده به عنوان نزدیک‌ترین تصویر چهره به تصویر آزمون.

در این تحقیق، چندین معیار فاصله دیگر نیز بررسی شده‌اند. نتایج بررسی‌ها در جدول ۱ نشان داده شده است. از بین معیارهای فاصله در روش پیشنهادی، دو معیار ذکر شده در جدول ۱ بهترین عملکرد را داشته‌اند. مقایسه‌ی این معیارها روی پایگاه داده با تعداد ۲۰۰ چهره انجام گرفته است.

جدول۱: مقایسه معیار فاصله اقلیدسی با دو معیار دیگر در روش پیشنهادی

معیار فاصله	میزان کاهش نرخ بازیابی نسبت به فاصله اقلیدسی	میزان افزایش زمان بازیابی نسبت به فاصله اقلیدسی
Canborra	٪۱	۱۳۲/۷۵ms
Jensen	٪۵	۸۱۲/۵ms

۳- آزمایش‌ها

جهت ارزیابی الگوریتم پیشنهادی، از پایگاه داده FERET استفاده شده است. این پایگاه داده شامل ۹۹۰ فرد متمایز و چندین چهره برای هر فرد است. در این مقاله، از زیرمجموعه‌های fa و fb این پایگاه داده استفاده شده است. این دو زیرمجموعه شامل تصاویر چهره از روبرو هستند. fa شامل تصاویر ساده و fb تصاویر متناظر با fa و همراه با ایجاد تغییرات حالت چهره است.

در الگوریتم پیشنهادی، از زیرمجموعه fa به عنوان تصاویر پایگاه داده و از زیرمجموعه fb به عنوان تصاویر آزمون استفاده شده است. در این بخش، روش پیشنهادی در دو مرحله ارزیابی می‌شود. ابتدا نتایج حاصل از الگوریتم پیشنهادی، با روش‌های اخیر مقایسه می‌شود. سپس توانایی الگوریتم پیشنهادی در بازیابی چهره، بر اساس تغییرات حجم پایگاه داده ارزیابی می‌شود.

هدف اصلی الگوریتم ارائه شده، بازشناسی چهره در حجم زیاد پایگاه داده و وجود تنها یک تصویر چهره به ازای هر شخص است. از بین روش‌های اخیر در بازشناسی چهره، دو روش از پایگاه داده‌ی حجیم‌تری نسبت به سایر روش‌ها استفاده کرده‌اند که نتایج قابل قبول‌تری نسبت به سایرین حاصل کرده‌اند[۶، ۱۰]. اولین روش، الگوریتم SIFT در [۱۰] است. این روش، از پایگاه داده FERET با ۲۰۰ چهره انتخابی استفاده و نتایج خود را ثبت کرده است. الگوریتم دوم، الگوریتم LBP در [۶] است. در این الگوریتم، ۱۶۰ چهره از پایگاه داده FERET انتخاب شده است. در این روش، از هر چهره ۴ تصویر در اختیار بوده است که ۱ تصویر در پایگاه داده جهت بازیابی و ۳ تصویر جهت آزمون مورد استفاده قرار گرفته‌اند. همچنین روش پیشنهادی با روش‌های دیگر که دارای حجم پایگاه داده کمتر هستند مقایسه شده است. در جدول ۲ نتایج ارزیابی الگوریتم ارائه شده با روش‌های اخیر نشان داده شده است.

جدول۲: مقایسه کارآیی الگوریتم‌های مختلف با روش پیشنهادی

الگوریتم	تعداد افراد پایگاه داده	تعداد تصویر از هر فرد	پایگاه داده	نرخ بازیابی
DCP[19]	۱۲۶	۲۶	AR	۹۴/۶۴
Hough[14]	۳۵	۲۰	FERET	٪۸۰/۸۱
Gabor[15]	۳۵	۲۰	FERET	٪۸۳/۳
SIFT[10]	۲۰۰	۱	FERET	٪۹۷
LBP[6]	۱۶۰	۴	FERET	٪۹۷
روش پیشنهادی	۲۰۰	۱	FERET	٪۹۹

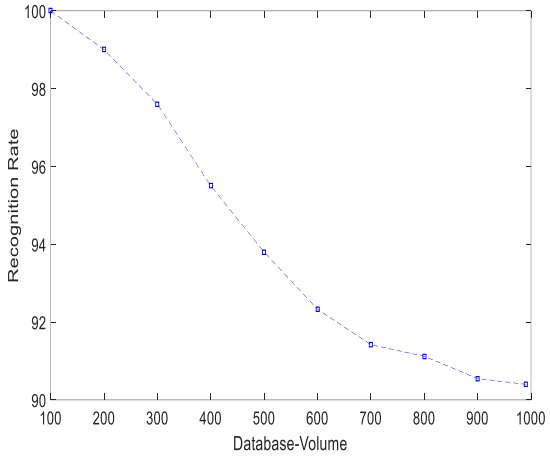
مزیت اصلی الگوریتم پیشنهادی، استفاده از تمام ناحیه چهره جهت استخراج ویژگی است. این در حالی است که به دلیل افزایش ابعاد ویژگی‌ها، سرعت بازیابی بسیار مناسب است. روش پیشنهادی با پیمایش کل ناحیه چهره و استخراج ویژگی‌های مناسب، قابلیت تفکیک بین فضای ویژگی افراد مختلف در حجم زیاد را فراهم آورده است. اکثر روش‌های بازشناسی چهره، جهت کاهش زمان بازیابی، از الگوریتم‌های مدل‌گرا استفاده می‌کنند که در نتیجه، به دلیل مد نظر قرار دادن نواحی خاص از چهره، نرخ بازشناسی کاهش می‌یابد. برخی روش‌ها برای حل مشکل کمبود ویژگی‌ها از چندین تصویر به ازای هر شخص استفاده می‌کنند که خود باعث افزایش حجم پایگاه داده و سخت شدن کاربری می‌شود. در الگوریتم ارائه شده، به ازای هر شخص تنها یک چهره در پایگاه داده وجود دارد و سرعت بازیابی آن، بسیار بالا است. در جدول ۳ نتایج الگوریتم پیشنهادی در حجم‌های متفاوت پایگاه داده، آورده شده است.

جدول۳: نرخ و سرعت بازیابی الگوریتم ارائه شده در حجم‌های متفاوت پایگاه داده

نرخ بازیابی	سرعت بازیابی(ms)	حجم پایگاه داده
٪۱۰۰	۹۹/۵	۱۰۰
٪۹۹	۱۸۶	۲۰۰
٪۹۷/۶۶	۲۵۶	۳۰۰
٪۹۵/۵	۳۴۳	۴۰۰
٪۹۳/۸	۴۳۸	۵۰۰
٪۹۲/۳۳	۵۲۲	۶۰۰
٪۹۱/۴۳	۶۳۶	۷۰۰
٪۹۱/۱۲	۷۴۰	۸۰۰
٪۹۰/۵۵	۷۹۰	۹۰۰
٪۹۰/۴	۸۶۰	۹۹۰

در این بررسی، هر بار به صورت تصادفی تعداد مختلفی تصویر از پایگاه داده انتخاب می‌شوند. روش‌های اخیر، تعداد افراد را بیشتر از ۲۰۰ لحاظ نکرده‌اند. با توجه به آزمایش‌های انجام شده با افزایش تعداد اشخاص، بازدهی این روش‌ها به شدت افت می‌کند.

با توجه به تاکید این مقاله بر حجم نمونه‌ها، چنانچه در شکل ۸ نشان داده شده است، شتاب کاهش نرخ بازیابی تصویر چهره با افزایش تعداد تصاویر پایگاه داده، بسیار کم است که کارایی مناسب روش پیشنهادی در حجم زیاد نمونه‌ها را نشان می‌دهد.



شکل ۸: شتاب کاهش نرخ بازیابی با افزایش حجم پایگاه داده

۴- نتیجه

در این مقاله روشی نوین برای بازشناسی چهره با استفاده از ویژگی‌های مکان- فرکانس ارائه شده است. در این روش، ویژگی‌های مناسبی از تصویر چهره در حوزه فرکانس و با محاسبه تغییرات مولفه‌های فرکانسی، استخراج شده است. همچنین در این مقاله از فیلتری در حوزه فرکانس، جهت کاهش ابعاد ویژگی و بهبود نرخ بازیابی استفاده شده است. روش استخراج ویژگی، جدید بوده و تا کنون از آن جهت بازیابی چهره استفاده نشده است. در نهایت با استفاده از معیار فاصله اقلیدسی، نزدیک‌ترین تصویر پایگاه داده بازشناسی شده است. الگوریتم پیشنهادی، در کاربردهایی با حجم پایگاه داده انبوه، نسبت به سایر الگوریتم‌ها بهتر ایفای نقش می‌کند. همچنین این الگوریتم در پایگاه داده‌هایی با تنها یک تصویر به ازای هر شخص، بازدهی بسیار مناسبی دارد. در این مقاله، نتایج حاصل با روش‌های پیشین مورد مقایسه قرار گرفته و نشان داده شده است که نرخ بازیابی از روش‌های قبلی بیشتر بوده و به ٪۹۹ ارتقا پیدا کرده است. همچنین با افزایش حجم پایگاه داده تا ۹۹۰ تصویر چهره، نرخ بازیابی مناسبی حاصل گردیده و نشان داده شده است که شتاب کاهش نرخ بازیابی با افزایش حجم پایگاه داده بسیار کم است. این خاصیت بیانگر نرخ بازیابی مناسب در افزایش چندین برابری حجم پایگاه داده، است.

مراجع

[1] Stan Z. Li and Anil K.. Jain, , *Handbook of Face Recognition*, Edition 2, Springer Publishing, 2011.

[۲] شفیع‌پور یوردشاهی، سیدعربی و آقاگل‌زاده، « بازشناسی چهره بر اساس ظاهر در رشته‌های ویدئویی»، مجله مهندسی برق دانشگاه تبریز، جلد ۴۵، شماره ۲، صفحات ۷۵-۸۳، تابستان ۱۳۹۴.

[3] L. Lenc and P. Kral, Face Recognition under Real-world Conditions, ICAART(2), pp. 250-256, 2013.

[۴] صادقی، آیت‌اللّهی و راعی، « بازشناسی حالت چهره با استفاده از نرمالیزاسیون هندسی و تبدیل موجک مختلط Dual-Tree»، مجله مهندسی برق دانشگاه تبریز، جلد ۴۵، شماره ۳، صفحات ۷۹-۸۷، پاییز ۱۳۹۴.

[5] M. Turk and A. Pentland, “Eigenfaces for recognition", Journal of Cognitive Neuroscience, Vol. 3, no. 1, pp.71-86, 1991.

[6] T. Ahonen, A. Hadid and M. Pietikainen, "Face recognition with local binary patterns", Computer Vision - ECCV, Vol. 302, pp.469-481, 2004.

[7] S. Liao, X. Zhu, Z. Lei, L. Zhang and Stan Z. Li, "Learning multi-scale block local binary patterns for face recognition", Advances in Biometrics, Vol. 4642, pp. 828-837, 2007.

[8] D. H. Wang and P. Conilione, "Machine learning approach for face image retrieval", Neural Computing and Applications, Vol. 21, no. 4, pp.683-694, 2012.

[9] M. J. Er, W. Chen and S. Wu, "High-speed face recognition based on discrete cosine transform and RBF neural networks", IEEE Transactions on Neural Networks , Vol. 16, no. 3, pp. 679-691, 2005.

[10] J. Luo, Y. Ma, E. Takikawa , S. Lao, M. Kawade and B. Lu, "Person-specific SIFT features for face recognition", IEEE International Conference on Acoustics, Speech and Signal Processing, Vol. 2, pp. II-593, 2007.

[11] L. Lenc and P. Král , "A combined SIFT/SURF descriptor for automatic face recognition", Sixth International Conference on Machine Vision, Vol. 9067, pp. 90672C-90672C, 2013.

[12] M. Zhou , H. Wei and S. Maybank, "Gabor wavelets and AdaBoost in feature selection for face verification", Workshop in application of computer vision, pp.101-109, 2006.

[13] H. Cho, R. Roberts, B. Jung, O. Choi and S. Moon, "An efficient hybrid face recognition algorithm using PCA and GABOR wavelets", International Journal of Advanced Robotic Systems, Vol. 11, p.59, 2014.

[14] R. Varun, Y. VivekanandKini, K. Manikantan and S. Ramachandran, "Face recognition using hough transform based feature extraction", Procedia Computer Science,Vol. 46, pp.1491-1500, 2015.

[15] T. M. Abhishree, J. Latha, K. Manikantan and S. Ramachandran, "Face recognition using gabor filter based feature extraction with anisotropic diffusion as a pre-processing technique", Procedia Computer Science, Vol. 45, pp. 312-321, 2015.

[16] C. Ma, J. Y. Jung, S. W. Kim and J. S. Ko, "Random projection-based partial feature extraction for robust face recognition", Neurocomputing, Vol. 149, pp. 1232-1244, 2015.

[17] S. Wang and P.Liu, "A new feature extraction method based on the information fusion of entropy matrix and covariance matrix and its application in face recognition", entropy, Vol. 17, no. 7, pp. 4664-4683, 2015.

[18] X. Xie and K. M. Lam, "Face recognition using elastic local reconstruction based on a single face image", Pattern Recognition, Vol. 41, no. 1, pp. 406–417, 2008.

[19] Y. Gao, Y. Qi, "Robust visual similarityretrieval in single model face databases", Pattern Recognition, Vol. 38, no. 7, pp. 1009-1020, 2005.

[20] J. Haddadnia, K. Faez and M. Ahmadi, "A fuzzy hybrid learning algorithm for radial basis function neural network with application in human face recognition", Pattern Recognition, Vol. 36, no. 5, pp. 1187–1202, 2003.

[21] J. Lu, X. Yuan and T. Yahagi, "A method of face recognition based on fuzzy c-means clustering and associated Sub-NNs", IEEE Transactions On Neural Networks, Vol. 18, no. 1, pp. 150 – 160, 2007.

[22] P. Viola and M. Jones,"Rapid object detection using a boosted cascade of simple features", In Proc. of IEEE Conference on Computer Vision and Pattern Recognition, Vol. 1, pp. I-I, 2001.

[23] K. K. Paliwal and L. D. Alsteris, "On The usefulness of STFT phase spectrom in human listening tests", Speech Communication, Vol. 45, no. 2, pp. 153-170, 2004.

زیرنویس‌ها

[1] Holistic matching methods

[2] Feature-based (structural) matching methods

[3] Hybrid methods

[4] Eigenface

[5] Chen

[6] Fisher

[7] Radial basis function

[8] Scale-invariant feature transform

[9] Speeded Up Robust Features

[10] Zhou

[11] Adaboost

[12] Principal Component Analysis

[13] Fuzzy c-means clustering

[14] Decibel

[15] Short time fourier transform